

## Causal discounting in the presence of a stronger cue is due to bias

JEFFREY P. LAUX

*University of Texas, Austin, Texas*

KELLY M. GOEDERT

*Seton Hall University, South Orange, New Jersey*

AND

ARTHUR B. MARKMAN

*University of Texas, Austin, Texas*

People use information about the covariation between a putative cause and an outcome to determine whether a causal relationship obtains. When there are two candidate causes and one is more strongly related to the effect than is the other, the influence of the second is underestimated. This phenomenon is called *causal discounting*. In two experiments, we adapted paradigms for studying causal learning in order to apply signal detection analysis to this phenomenon. We investigated whether the presence of a stronger alternative makes the task more difficult (indexed by differences in  $d'$ ) or whether people change the standard by which they assess causality (measured by  $\beta$ ). Our results indicate that the effect is due to bias.

Humans can use knowledge of covariation to predict events and to infer their underlying causes (Cheng, 1997). Although research has demonstrated a number of systematic phenomena in covariation and causal judgment, it is unclear whether these effects occur during the learning or the decision process. Here, we use signal detection theory (SDT) to tease apart these alternatives for one phenomenon: causal discounting.

Causal discounting is a *cue interaction* effect, in which someone judges a moderately effective cause as less effective when it is learned about in the presence of a highly effective alternative (e.g., Baker, Mercier, Vallee-Tourangeau, Frank, & Pan, 1993; Goedert & Spellman, 2005). For example, a person taking a steroid and an antihistamine for allergies may believe a 50%-effective antihistamine is less effective, if used with a 90%-effective steroid. Another type of cue interaction arises when the occurrence of two cues is confounded and participants control for the alternative cue when judging the target (Spellman, 1996). Here, we focus on the case in which participants devalue a target cause in the presence of an alternative and the causes are *unconfounded*—that is, causal discounting proper (Goedert, Harsch, & Spellman, 2005).

Although cue interaction phenomena are reliably observed in both causal judgment and prediction, it is debated whether these phenomena reflect learning processes or decision processes (Stout & Miller, 2007). For instance, in our example above, the perceived effectiveness of the antihista-

mine may be lower because less is learned about it when the more effective steroid is present. Alternatively, the highly effective steroid may bias the judgment process by “raising the bar” for effectiveness. Signal detection analyses allow one to determine whether changes in performance reflect changes in the participant’s sensitivity—that is, the ability to detect contingency between the cause and outcome—or changes in the participant’s decision criterion.

SDT (Macmillan & Creelman, 2005) is a data-analytic tool that disentangles a person’s sensitivity to detect a stimulus ( $d'$ ) from that person’s bias to say “yes” ( $\beta$ ). These latent variables are calculated from two components of participants’ responses: The hit rate ( $h$ ) is the proportion of trials on which participants say “yes” when the candidate is causal. The false alarm rate ( $fa$ ) is the proportion of trials on which participants say “yes” when the candidate is noncausal.

SDT has successfully differentiated learning from decision processes in memory research. For example, training in mnemonic techniques affects sensitivity, leaving bias unchanged (McNicol & Ryder, 1971). Conversely, changing the payoff structure for correct responses affects bias, not sensitivity (Healy & Kubovy, 1978). A brief sketch will illustrate these ideas. Imagine a word list in which people typically recognize 62.5% of the old words ( $h$ ) but also say that they remember 47.5% of the lures ( $fa$ ). One group is taught a new study strategy for remembering words. At test, their performance is much improved, relative to the

---

J. P. Laux, laux@mail.utexas.edu

---

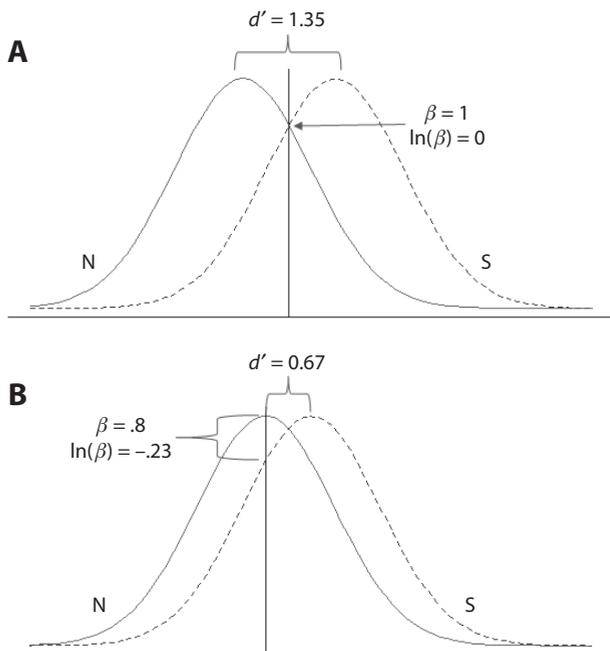
standard:  $h = 75\%$  and  $fa = 25\%$  (Figure 1A). Another group is sternly warned not to miss any words. Their rates differ:  $h = 75\%$  and  $fa = 50\%$  (Figure 1B). It is clear that the mnemonic increases retention, but the warning only makes participants more likely to say “yes.”

SDT makes some processing assumptions. First, there is ambient noise in people’s representational systems. Over time, the value represented will vary around a mean (Figure 1, N distributions). When a signal occurs (e.g., an old word), the value increases by an amount proportional to the signal’s strength. Thus, the signal distribution is similar but shifted (Figure 1, S distributions). The distance between the distributions is  $d'$ : how easily they are differentiated. It is calculated by passing the rates through an inverse cumulative distribution function (yielding  $z$  scores) and finding the difference:

$$d' = z(h) - z(fa). \tag{1}$$

A second assumption is that a person will say “yes” if the current value is above some threshold (Figure 1, vertical lines). This criterion could be anywhere, but optimal accuracy is achieved where the distributions cross (Figure 1A). Such a person is unbiased: At that point, the likelihood that the value is a signal equals the likelihood that it is noise. However, some people may be overeager (Figure 1B) or excessively hesitant to say “yes.” The ratio of the likelihoods at the criterion is  $\beta$ , the relative amount of evidence required to say “yes”:

$$\beta = \frac{f[z(h)]}{f[z(fa)]}, \tag{2}$$



**Figure 1.** Example noise (N, solid) and signal (S, dashed) distributions, as they are constituted in signal detection theory, for an unbiased group that distinguishes signals and noise very easily (A), and an overeager group that can differentiate only moderately well (B).

where  $f(x)$  is the normal distribution’s likelihood function.

Because  $\beta$  ranges from 0 to  $+\infty$ , with  $\beta = 1$  as unbiased, its distribution necessarily is highly skewed and is typically transformed via the natural logarithm. We follow this policy. Thus,  $\ln(\beta)$  is positive for those participants hesitant to say “yes” and negative for overeager participants.<sup>1</sup>

Our brief sketch above also suggests a quasi-SDT analysis that avoids SDT’s assumptions: The easier it is to differentiate between true causes and uncorrelated candidates, the greater the difference should be between  $h$  and  $fa$ . In addition, it is clear that the overall tendency to say “yes” is indexed by the mean of  $h$  and  $fa$ . These measures are quite similar to  $d'$  and  $\beta$  and can be used when SDT’s requirements are not met.

Researchers have begun to apply SDT to causal induction, primarily examining situations in which participants evaluate a single candidate cause (Allan, Siegel, & Tangen, 2005; Perales, Catena, Shanks, & Gonzalez, 2005). For example, Allan et al. (2005) found that participants’ sensitivity increased as *contingency* increased but that bias varied with the base rate of the outcome. Our present experiments of causal discounting are among the first to apply SDT to the case in which there are two candidate causes of a common outcome. (Recently, SDT has begun to be applied to the blocking effect as well; Siegel, Allan, Hannah, & Crump, 2009.)<sup>2</sup>

To apply SDT to the case of discounting, we adopted the streamed-trial technique developed by Allan, Hannah, Crump, and Siegel (2008). In a streamed trial, participants view a large number of events in one unbroken stream before responding. Each trial contains complete contingency information, and these contingencies can change from trial to trial. Importantly, participants’ responses on each trial can be objectively correct or incorrect (i.e., hits and false alarms). Crump, Hannah, Allan, and Hord (2007) have validated the streamed-trial method by replicating standard phenomena of contingency learning.

We extended the streamed-trial technique to the two-cause case. Like standard SDT, in which the stimulus is either present or absent, participants saw one of two trial types: target causal or target noncausal. We independently varied the strength of the alternative cause. An example of a streamed trial appears in Figure 2. At the end of each streamed trial, participants responded whether the target cause increased the probability of the effect. On some proportion of trials on which the target is causal, the participants will say “yes,” but they may also say “yes” when the target is noncausal (producing  $h$  and  $fa$ , respectively).

## METHOD

Experiments 1 and 2 differed only in their participants, cover story, and stimuli.

### Participants

Undergraduates at the University of Texas, Austin (Experiment 1, 90, with 45 in each condition; Experiment 2, 195, with 97 in the strong alternative [SA] condition and 98 in the weak alternative [WA] condition) participated for course credit.

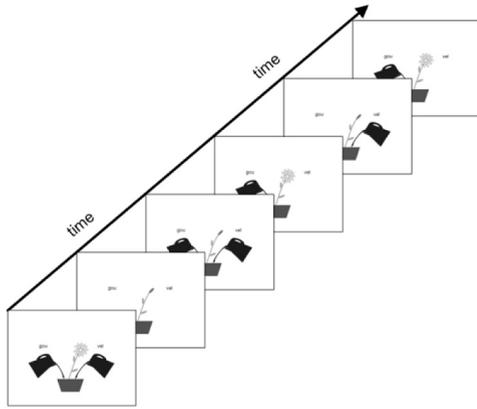


Figure 2. Example of a streamed trial from Experiment 1.

**Design and Contingency Structures**

We employed a single between-subjects factor (alternative strength: strong vs. weak). We measured contingency as  $\Delta P$  (Allan, 1980), the change in probability of the outcome, given the presence of the candidate cause [ $P(O|C)$ ] from its absence [ $P(O|\sim C)$ ]:

$$\Delta P = P(O|C) - P(O|\sim C). \tag{3}$$

If the causes are unconfounded,  $\Delta P$  can be applied to multicausal situations by using marginal frequencies instead of cell frequencies (i.e., collapsing over the other causes).

In the SA condition,  $\Delta P$  of the alternative was .33, and in the WA condition,  $\Delta P_A = 0$ . Because SDT requires equal numbers of trials on which the answer is objectively “yes” and “no,” we devised two contingency structures for each condition: one in which the target was causal ( $\Delta P_T = .22$ ) and one in which it was not ( $\Delta P_T = 0$ ). Thus, the experiments employed four contingency structures (Figure 3), with  $\Delta P_A$  equivalent within conditions but differing between them and  $\Delta P_T$  differing among trials within conditions.

**Cover Story and Stimuli**

In Experiment 1, the participants determined whether liquids made it more likely that a flower would bloom. On each trial, the participants viewed multiple events depicting a plant blooming or not, with zero, one, or two watering cans pouring liquid onto the plant. Above each can was displayed a pronounceable three-letter nonword identifying the liquid (Rastle, Harrington, & Coltheart, 2002). Each trial used unique names, reinforcing the independence of the trials.

In Experiment 2, the participants were medical researchers determining whether medicines made it more likely that patients got well. The stimuli were drawings of a smiley face (recovery) or a sickly green face (death) with a pill on one, both, or neither side. Above each pill was written a pronounceable three-letter nonword that stood for its chemical name and that varied from trial to trial.

**Procedure**

The participants were tested individually on computers. After reading instructions and viewing sample slides, they viewed 72 streamed trials. Each streamed trial consisted of 36 events and contained one of the contingency structures depicted in Figure 3. Each event was displayed for 550 msec, with a 100-msec interstimulus interval; each trial lasted 23.4 sec. Each participant saw a mixture of only two types of streamed trials constituting the contingency structures (target causal and noncausal) for their condition.

The participants initiated each trial by hitting “Enter.” At the end of each streamed trial, the participants responded whether they thought that the probability of the outcome (i.e., blooming or recovering) was increased by the cause on the left (i.e., liquid or drug) and

then by the cause on the right. The target’s side (left or right) was counterbalanced within participants.

**RESULTS**

Differentiating  $\Delta P$ s of 0 and .22 is very difficult, and a number of our participants failed to discriminate between the target causal and noncausal trials. Twelve participants in Experiment 1 (SA, 3; WA, 9) and 44 in Experiment 2 (SA, 23; WA, 21) had  $d' \leq 0$ . Chi-square analyses suggested that the inability to discriminate trial types did not vary with condition [ $\chi^2(1) = 3.6, p = .06$ , and  $\chi^2(1) = 0.3, p = .58$ , for Experiments 1 and 2, respectively]. When  $d' \leq 0$ , there is no signal distribution, and  $\beta$  is meaningless. We thought it important to conduct analyses for sensitivity and bias over the same participants. Thus, for both experiments, we report two sets of analyses: SDT analyses for the participants whose  $d' > 0$  (i.e., the participants who were able to discriminate the trial types), but also the quasi-SDT analyses for all the participants. Throughout, we adopt  $\alpha = .05$ .

There was no evidence of learning or fatigue across trials. We scored for accuracy by assigning 1 to every trial on which the participants judged the target correctly and 0 if incorrect. Collapsing over both experiments, correlations between trial number and accuracy were normally distributed, with  $M = -.01, SD = .12$ . Consistent with our instructions that trials were unrelated, the participants’ responses on each trial were independent of previous responses. Lag 1 autocorrelations for the right and left cues were normally distributed, with  $M = -.01, SD = .14$ .

Our key results are displayed in Table 1 (“yes” rates) and Table 2 (SDT parameters). Both tables report descriptive statistics for the filtered data.

		Weak Alternative		Strong Alternative	
		A	~A	A	~A
Target Causal	T	6/9	6/9	7/9	4/9
	~T	4/9	4/9	5/9	2/9
		.67	.44	.61	.39
		.56	.56	.67	.33
		$\Delta P_T = .22$		$\Delta P_T = .22$	
		$\Delta P_A = 0$		$\Delta P_A = .33$	
		A	~A	A	~A
Target Not Causal	T	5/9	5/9	6/9	3/9
	~T	5/9	5/9	6/9	3/9
		.56	.56	.67	.33
		.56	.56	.67	.33
		$\Delta P_T = 0$		$\Delta P_T = 0$	
		$\Delta P_A = 0$		$\Delta P_A = .33$	

Figure 3. Contingency structures used in Experiments 1 and 2. In each cell, the denominator represents the number of events with that combination of cues; the numerator indicates the number of times the positive outcome occurred.

**Table 1**  
Means (and Standard Deviations) of “Yes” Rates

	Alternative							
	Experiment 1				Experiment 2			
	Strong		Weak		Strong		Weak	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Target (causal)	.43	.19	.59	.16	.48	.16	.67	.14
Target (not causal)	.27	.16	.37	.17	.28	.14	.47	.13
Alternative (target causal)	.54	.19	.35	.16	.61	.14	.43	.19
Alternative (target not causal)	.60	.21	.41	.16	.69	.14	.50	.14

Note—The “yes” rates are mean *h* and *fa*. Mean *d'* and  $\beta$  cannot be calculated from mean “yes” rates, because they must be calculated for each individual first (which involves nonlinear transformations) and then averaged.

**Experiment 1**

By common standards, causal discounting occurred. The participants thought that the causal target was causal more often in the WA condition than in the SA condition [ $t(76) = 3.8, d = 0.87$ ]. Interestingly, this was also true of the participants’ inferences about the noncausal target: They said “yes” more in the WA condition than in the SA condition [ $t(76) = 2.9, d = 0.65$ ].

There was no effect on sensitivity [ $t(76) = 0.8, p = .41, d = 0.19$ ]. There was, however, a marginal effect on bias. The SA group was less willing to infer that the target was causal than was the WA group [ $t(76) = 1.8, p = .08, d = 0.41$ ].

Although our primary manipulation was between subjects, we can examine the participants’ tendencies to say “yes” to the *alternative* when the target was causal versus noncausal. These rates differ: In both the SA condition [ $t(41) = 3.0, d = 0.46$ ] and the WA condition [ $t(35) = 2.6, d = 0.44$ ], the participants were more likely to call the alternative causal when the *target* was not causal. Although SDT analyses cannot be applied here, if these within-subjects differences in the perceived causal strength of the alternative also reflect a shift in the participants’ decision criterion (as was the case for the between-subjects analysis of responding to the target), these results suggest that the participants shifted their criterion on every trial, judging each cue relative to the other.

We also ran quasi-SDT analyses using all the participants. The mean difference between *h* and *fa* (i.e., how much more often they said “yes” to the causal than to the noncausal target) did not differ between the SA ( $M = .15$ ) and the WA ( $M = .17$ ) conditions [ $t(88) = 0.6, p = .52, d = 0.13$ ]. However, the average of *h* and *fa* (i.e., how often they said “yes” to targets overall) did vary, with the SA group ( $M = .35$ ) being substantially less willing to affirm the target than was the WA group ( $M = .47$ ) [ $t(88) = 3.6, d = 0.76$ ].

**Experiment 2**

Causal discounting was observed. The participants responded “yes” to the target more in the WA condition than in the SA condition, both when it was causal [ $t(149) = 8.2, d = 1.34$ ] and when it was not [ $t(149) = 8.6, d = 1.39$ ].

Again, there was no effect on sensitivity [ $t(149) = 0.4, p = .69, d = 0.06$ ]. Conversely, the data on bias in Experiment 2 are clear. As can be seen in Table 2, the participants

in the SA group were hesitant to conclude that the target was causal. Interestingly, the WA group was also biased, albeit less and in the other direction: They were overeager to say “yes.” As this suggests, the conditions differed substantially on bias [ $t(149) = 8.5, d = 1.25$ ].

In addition, consistent with Experiment 1, the participants found the alternative causal more often when the target was noncausal in both the SA [ $t(73) = 5.0, d = 0.58$ ] and the WA [ $t(76) = 3.8, d = 0.43$ ] conditions. In combination with the SDT analyses above, it appears that the participants changed the standard by which they determined causality trial by trial.

These findings are further supported by analyses including all the participants. There was no difference in the gap between *h* and *fa* for the SA ( $M = .15$ ) or the WA ( $M = .16$ ) [ $t(193) = 0.4, p = .66, d = 0.06$ ], but there was a huge difference in the SA’s ( $M = .38$ ) and WA’s ( $M = .56$ ) mean rates of positive responses [ $t(193) = 10.9, d = 1.39$ ].

**DISCUSSION**

We demonstrated discounting in two experiments using Allan et al.’s (2008) streamed-trial procedure: The participants were less likely to say that the target was causal when it was learned in the presence of a more contingent alternative. Critically, we also demonstrated that discounting is due not to impaired ability to detect covariation but, rather, to changes in participants’ response criterion. The participants’ ability to discriminate between causal and noncausal trials did not vary with the strength of the alter-

**Table 2**  
Means (and Standard Deviations) of Signal Detection Theory Parameters

Alternative	<i>d'</i>		$\ln(\beta)$	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experiment 1				
Strong	0.53	0.32	0.22	0.46
Weak	0.59	0.37	0.07	0.25
Cohen’s <i>d</i>	0.19		0.41	
Experiment 2				
Strong	0.60	0.38	0.24	0.35
Weak	0.57	0.35	-0.14	0.25
Cohen’s <i>d</i>	0.06		1.25	

native. However, learning about the target in the presence of the strong alternative made the participants hesitant to call the target causal.

### Potential Limitations

SDT requires that participants make both “yes” and “no” responses to the target in all trial types. Considerable effort was required to find a set of experimental parameters (contingency structures, presentation times, etc.) that yielded viable data. This led us to the paradigm we employed. Among other things, it meant that we had contingency structures with a relatively small difference between the causal and noncausal target  $\Delta P$ s, which makes contingency discrimination difficult (Allan et al., 2005). Thus, there were many nonlearners. Despite these aspects of the procedure, we have reason to believe the generalizability of our results. (1) More than 80% of the participants (229 of 285 across both experiments) did discriminate between the target causal and target noncausal trials. (2) We replicated causal discounting as demonstrated with other contingency structures and methods of responding.

In addition, the contingency structures we used differed trivially in the probability of the outcome [SA,  $P(O) = .50$ ; WA,  $P(O) = .56$ ]. Because responding increases as the base rate of the outcome increases (i.e., the *outcome density effect*; Allan et al., 2005), we examined whether differences in outcome density would explain our results. They did not. Using Allan et al.’s (2005) data for responding after 40 events, we fit a regression model and predicted the expected increase, using our contingency structures. The regression model suggests that our responding should increase by .15 SDs for the causal and .08 SDs for the noncausal targets. Thus, although differences in  $P(O)$  could have inflated our effects slightly, they cannot account for our results. Furthermore, within each group and experiment, responding to the alternative varied significantly with the  $\Delta P$  of its “alternative” (i.e., the target), even though  $P(O)$  was the same for those contrasts.

Researchers familiar with SDT in the context of psychophysics may find a between-subjects design unusual, but between-subjects designs are commonly and successfully used when SDT is applied in memory research. Although a within-subjects design might seem to increase our power, with only 72 trials, it would mean that participants would see only 18 of each trial type; thus, it would greatly increase measurement error. Moreover, our analyses of responding to the alternative replicates discounting within subjects: The participants responded “yes” less on trials on which the target was causal than on trials on which it was noncausal.

Finally, we acknowledge that the streamed-trial procedure is a relative newcomer in causal-reasoning research. Further research exploring the relations between this task and other methods used to study causal induction would be helpful.

### Theoretical Implications

Traditionally, theories of causal induction from covariation have been either associative or statistical in nature. Associative theories assume that an association is built by

translating cue co-occurrences into a summary value as they are experienced (e.g., Rescorla & Wagner, 1972). In contrast, statistical theories posit that people store memory traces of past events and run computations that approximate statistical analyses over these observed frequencies (e.g., Schustack & Sternberg, 1981). Current theorizing allows for multiple processes (e.g., Cheng, 1997; Perales et al., 2005; Stout & Miller, 2007).

Research supports the hypothesis that causal inference from covariation involves processes beyond the prediction of events. For example, causal judgments might be expected to rely on frequency estimates; yet causal judgments are subject to cue interaction effects, but frequency estimates are not (Price & Yates, 1995). Causal judgments also vary with the structure of the scenario (*common cause vs. common effect*), but predictions do not (Tangen & Allan, 2004). Thus, unique effects exist for causal judgments per se.

If sensitivity and bias are mapped onto two sets of processes, the simplest formula is sensitivity  $\rightarrow$  learning and bias  $\rightarrow$  judgment. Essentially, the more you learn about something, the better you are at identifying it accurately (i.e., as contingent or not), whereas once you have the evidence for covariation, whether to say “yes” is a judgment. From this perspective, our findings imply that discounting involves judgment rather than learning processes. The influence of judgment processes is consistent with other research. For example, participants presented with summary tables and asked whether they think that a cue is causal also demonstrate discounting (Goedert & Spellman, 2005). It is difficult to attribute that discounting to an associative-learning process, because the series of experiences necessary to build an association is lacking.

At first glance, our results do not appear to reinforce associationist accounts. Although numerous associative theories have been proposed, most posit differences in learning as the proximate cause of differences in behavior (Stout & Miller, 2007). It is impossible to state exactly how all such theories should map onto SDT’s parameters, but differences in learning without a separate judgment process seem most consistent with changes in sensitivity and not bias. This is the opposite of what we found. Of associative theories, only the comparator hypothesis (Stout & Miller, 2007) has a clearly separate judgment process that constitutes the locus of the theory’s effects.

Our results do not obviously favor statistical theories either, since they typically lack a clear mechanism leading people to assess causality more tentatively. Usually, they capture causal phenomena by positing that people run computations over subsets of their experiences (Cheng & Novick, 1992), weight various cue/outcome combinations differently, or adjust the final calculation. Of such approaches, perhaps the Power PC model (Cheng, 1997) does best. In our situation, it proposes that people would divide  $\Delta P$  by  $1 - P(O|\sim T, \sim A)$ . This operation, in conjunction with our contingency structures, would predict lower responding to the causal target in the SA condition than in the WA condition, but not to the noncausal target. Nevertheless, the spirit of the Power PC theory is that people accurately encode contingencies but make adjustments when required

to assess causality. This suggests that learning is equivalent but judgments differ, which is our interpretation.

Of course, our results do not prove conclusively that there are two sets of processes subserving causal induction. Another possibility is that people perceive covariation relatively, not absolutely. The perception of brightness, loudness, mass, and so forth, is known to be Weberian: People perceive them as relative to other magnitudes of the same type. Covariation could be perceived similarly. If so, a single process might produce results like ours by recognizing differences between contingent and noncontingent cues but treating them differently on the basis of the level of contingency of other causes in the same context. Unpacking such complex single-process accounts would require future theoretical investigations, but this demonstrates that a single process cannot simply be dismissed.

We tried to remain theoretically neutral here, because the full implications of these findings for causal theories require a different, and much longer, exposition. However, we suggest that our results may provide challenges for many theories.

## Conclusion

We applied SDT paradigms to the phenomenon of causal discounting to determine whether it is due to sensitivity or bias. Using the streamed-trial procedure, we demonstrated that discounting is due to changes in the criterion that participants use. The most straightforward interpretation is that causal discounting is a judgment phenomenon. Although further research is necessary to discriminate among more nuanced interpretations of these results, these findings are a first step in localizing this effect.

## AUTHOR NOTE

We thank Caren Rotello for gracious assistance with some of the intricacies of SDT. We very much appreciate Lorraine Allan for allowing us access to archival data on the outcome density effect and for inspiring this project. We also thank the members of the Similarity and Cognition Lab and the Cognition and Perception area of the University of Texas at Austin, as well as Michael Domjan, Bill Geisler, and Mariska Leunissen, for helpful comments and discussion. Correspondence concerning this article should be addressed to J. P. Laux, Department of Psychology, University of Texas, 1 University Station A8000, Austin, TX 78712 (e-mail: laux@mail.utexas.edu).

## REFERENCES

- ALLAN, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, *15*, 147-149.
- ALLAN, L. G., HANNAH, S. D., CRUMP, M. J. C., & SIEGEL, S. (2008). The psychophysics of contingency assessment. *Journal of Experimental Psychology: General*, *137*, 226-243.
- ALLAN, L. G., SIEGEL, S., & TANGEN, J. M. (2005). A signal detection analysis of contingency data. *Learning & Behavior*, *33*, 250-263.
- BAKER, A. G., MERCIER, P., VALLEE-TOURANGEAU, F., FRANK, R., & PAN, M. (1993). Selective associations and causality judgments: Presence of a strong causal factor may reduce judgments of a weaker one. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 414-432.
- CHENG, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, *104*, 367-405.
- CHENG, P. W., & NOVICK, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, *99*, 365-382.
- CRUMP, M. J. C., HANNAH, S. D., ALLAN, L. G., & HORD, L. K. (2007). Contingency judgments on the fly. *Quarterly Journal of Experimental Psychology*, *60*, 753-761.
- GOEDERT, K. M., HARSCH, J., & SPELLMAN, B. A. (2005). Discounting and conditionalization: Dissociable cognitive processes in human causal inference. *Psychological Science*, *16*, 590-595.
- GOEDERT, K. M., & SPELLMAN, B. A. (2005). Nonnormative discounting: There is more to cue interaction effects than controlling for alternative causes. *Learning & Behavior*, *33*, 197-210.
- HEALY, A. F., & KUBOVY, M. (1978). The effects of payoffs and prior probabilities on indices of performance and cutoff location in recognition memory. *Memory & Cognition*, *6*, 544-553.
- MACMILLAN, N. A. & CREELMAN, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.
- MCNICOL, D., & RYDER, L. A. (1971). Sensitivity and response bias effects in the learning of familiar and unfamiliar associations by rote or with a mnemonic. *Journal of Experimental Psychology*, *90*, 81-89.
- PERALES, J. C., CATENA, A., SHANKS, D. R., & GONZALEZ, J. A. (2005). Dissociation between judgments and outcome-expectancy measures in covariation learning: A signal detection theory approach. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *31*, 1105-1120.
- PRICE, P. C., & YATES, J. F. (1995). Associative and rule-based accounts of cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*, 1639-1655.
- RASTLE, K., HARRINGTON, J., & COLTHEART, M. (2002). 358,534 nonwords: The ARC nonword database. *Quarterly Journal of Experimental Psychology*, *55A*, 1339-1362.
- RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.
- SCHUSTACK, M. W., & STERNBERG, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General*, *110*, 101-120.
- SIEGEL, S., ALLAN, L. G., HANNAH, S. D., & CRUMP, M. J. C. (2009). Applying signal detection theory to contingency assessment. *Comparative Cognition & Behavior Reviews*, *4*, 116-134.
- SPELLMAN, B. A. (1996). Acting as intuitive scientists: Contingency judgments are made while controlling for alternative potential causes. *Psychological Science*, *7*, 337-342.
- STOUT, S. C., & MILLER, R. R. (2007). Sometimes competing retrieval (SOCR): A formalization of the comparator hypothesis. *Psychological Review*, *114*, 759-783.
- TANGEN, J. M., & ALLAN, L. G. (2004). Cue interaction and judgments of causality: Contributions of causal and associative processes. *Memory & Cognition*, *32*, 107-124.

## NOTES

- $d'$  also ranges from 0 to  $+\infty$ , but is not necessarily skewed and not typically transformed. For a full exposition of SDT, see Macmillan and Creelman (2005).
- We thank Tom Beckers for alerting us to this new contribution.

(Manuscript received April 8, 2009;  
revision accepted for publication November 15, 2009.)